

## СТАТИСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ СИНТЕЗИРУЕМЫХ МАССИВОВ БИЛЛИНГОВОЙ ИНФОРМАЦИИ

*Аннотация.* В статье рассмотрена структура биллинговой информации и ее статистические характеристики. Графиками продемонстрированы распределения, полученные в результате работы программного комплекса по сбору статистики. Проанализирована реалистичность синтезируемого массива биллинговой информации.

*Ключевые слова:* биллинг; статистические характеристики.

С целью формирования у студентов практических навыков по поиску и анализу взаимосвязей элементов современных телекоммуникационных систем разработан программный комплекс для генерации массивов биллинговой информации, состоящих из фонового биллинга и ситуационных задач (тестов). В докладе приводятся статистические характеристики синтезируемых массивов фонового биллинга.

Массив биллинговой информации представляет собой упорядоченную по временной метке совокупность записей [1, 2]. Основными значимыми для генерации полями являются AbonentIMSI (номер SIM-карты абонента), AbonentIMEI (тип и серийный номер устройства абонента), AbonentPhone (MSISDN номер телефона), LAC и CellID (поля, указывающие на привязку к базовой станции), BillTime (время события), CallDuration (продолжительность соединения) и BillingType (вид события).

Для синтеза массива соединений разработана статистическая модель поведения абонентов сотовой связи  $K_0 = \langle F_h, F_f, F_{\text{time}}, F_{\text{dur}}, F_n, F_a \rangle$ , учитывающая статистические характеристики биллинговой информации, описываемые функциями распределения:  $F_h$  — вероятности генерации типа события,  $F_f$  — длительности события,  $F_{\text{time}}$  — времени,  $F_{\text{dur}}$  — продолжительности,  $F_a$  — типа,  $F_n$  — количество соединений.

Данные распределения были собраны на основе личного биллинга при помощи почасового поиска, обработаны и представлены в формате графиков. Статистический анализ синтезируемого биллинга показывает, что его характеристики схожи с исходным массивом. В частности, распределение количества событий по дням недели демонстрирует всплеск активности в пятничный ве-

чер и падение на период выходных (рис. 1). Распределение по продолжительности звонков указывает на то, что большинство звонков длится менее 10 с.

График распределения количества событий по времени суток в биллинге, рассчитанного для города с полутора миллионным населением (рис. 2), показывает, что события обновления местоположения абонента (LocUpd) значительно превышают по количеству остальные типы событий (звонки, SMS и подключение к GPRS).

Одним из важнейших требований, которым должен удовлетворять синтезируемый массив биллинговой информации, является его реалистичность, то есть сходство его статических и динамических характеристик с параметрами биллинговой информации реальных сетей операторов сотовой связи. Предполагается, что массивы биллинговой информации представляют собой вид телетрафика и обладают свойством статистического самоподобия, измеряемого показателем Хэрста [3].

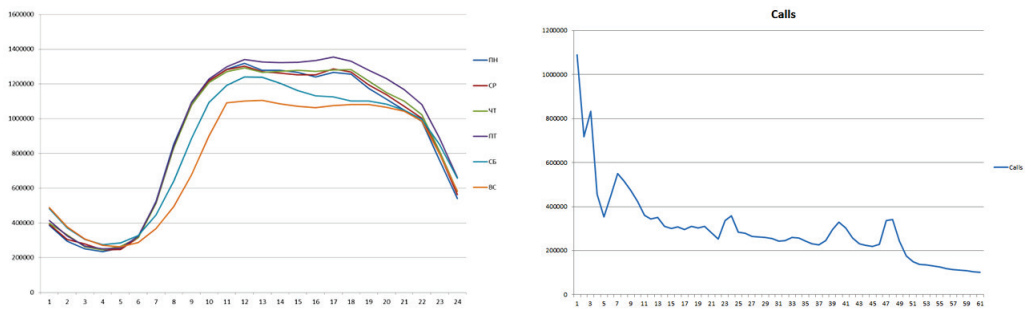


Рис. 1. Статистические распределения объема биллинга по дням недели и продолжительности звонков

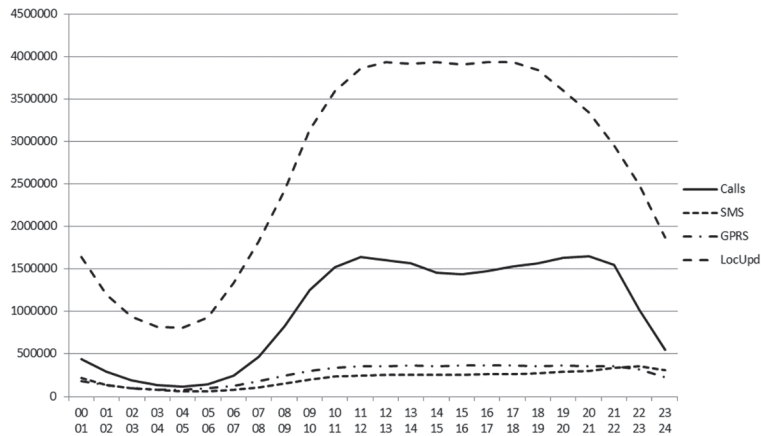


Рис. 2. График количественных отношений событий в биллинге в течение дня

Для определения данного свойства используются числовые характеристики распределений: отношение количества совершенных звонков к их длительности и временной промежуток между действиями абонентов в течение суток. Данные распределения в полной мере демонстрируют суть предлагаемой модели биллинговой информации, а также ее характеристики.

Полагается, что синтезированный массив биллинговой информации обладает свойством статистического самоподобия при значении показателя Хэрста больше 0,8.

Построенный на основе полученных статистических распределений биллинг по статистическим характеристикам идентичен биллинговой информации сетей операторов связи, обладает свойством самоподобия, измеряемым показателем Хэрста. Таким образом, можно говорить о статистической схожести синтезируемого биллинга с реальным и возможности его использования в составе учебных стендов при изучении поисково-аналитических комплексов.

### Список литературы

1. Семенищев И. А., Синадский А. Н., Синадский Н. И. Алгоритм формирования массива биллинговой информации на основе статистической модели поведения абонентов сотовой связи // XV Всерос. науч.-практ. конф. студентов, аспирантов и молодых ученых : сб. материалов. Курган : Курган. ГУ, 2016. С. 199–203.
2. Семенищев И. А., Синадский А. Н., Синадский Н. И. Статистические характеристики массива биллинговой информации при моделировании поведения абонентов сетей сотовой связи // 12-я Международ. молодеж. науч.-тех. конф : сб. материалов / под ред. А. А. Савочкина. Севастополь : Севастопол. гос. ун-т, 2016. С. 207.
3. Коллеров А. С. Критерий оценки качества синтезированного трафика на основе параметра Хэрста // Безопасность информационного пространства VI : сб. тр. межвуз. науч.-практ. конф. студентов, аспирантов и молодых ученых. Тюмень : Изд-во ТюмГУ, 2007. С. 32–34.